

Fig. 1: Amino acid sequence of the HBsAg a determinant of the different HBV genotypes as compared with the novel mutant HDB 05

A representative genome was used as the basis for each genotype and the aa sequence was deduced from the nucleotide sequence

A: X70 185; B: D00331; C: X01587; D: X72702, E: X75664; F: X75663; (Stuyver et al.; J. Gen. Virol. 81: 67-74 (2000); Norder et al.: J. Gen. Virol. 73: 3141-3145 (1992)

aa #	101	111	121	131	141	151	161	170
Genotyp								
A	QCMLPVCPLI	PGSITTTSTGP	CKTCTTPAQG	NSMFPSCCCT	KPTDGNCTCI	PIPSSWAF	AK	YLWEWASVRF
B	-----S-----	-----T-----	-----T-----	-----T-----	-----S-----	-----R-----	F	---G----
C	-----L--TS-----	-----I-----	-----R-----	-----Y-----	-----S-----	-----G-----	F	-----A----
D	-----S-----	-----R-----	-----R-----	-----M-----	-----L-----	-----S-----	---LG-----	-----A----
E	-----S-----	-----R-----	-----R-----	-----M-----	-----L-----	-----S-----	---LG-----	-----A----
F	-----S-----	-----R-----	-----R-----	-----M-----	-----L-----	-----S-----	---LG-----	-----A----
HDB 05	-----R---Q-----	-----L-----	-----V-----					
aa #	115	120	154	164				

The amino acid substitutions which differ from the wild-type adw HBV are printed in bold type

BEST AVAILABLE COPY

Fig. 2 Nucleotide sequence of the S gene of the known HBV adw wild type

encoding the HBV surface protein (surface antigen, HBsAg), and resulting amino acid sequence in the 3-letter and, especially, 1-letter codes (Coleman et al; WO 02/079217 A1)

Continuous numbering of nucleotides (nt) encoding the surface antigen (excl. pre S1 and pre S2 regions)

Continuous numbering of amino acids (aa)

	(aa)	(nt)
1	Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Val Leu Val Leu Gln Ala Gly Phe	60
61	M E N I T S G F L G P L L V L Q A G F F	20
121	TTG TTG ACA AGA ATC CTC ACA ATA CCG CAG AGT GTA GAC TCG TCG TCG ACT TCT CTC AAT	120
41	Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu Asp Ser Thr Thr Ser Leu Asn	40
181	L L T R I L T I P Q S L D S W T S L N	180
241	TTT CTA GGG GGA TCA CCG GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC	240
61	Phe Leu Gly Gly Ser Pro Val Cys Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His	60
121	F L Q G S P V C L G Q N S Q S P T S N H	120
181	TCA CCA ACC TCC TGT CCT CCA ATT TGT CCT GGT TAT CCG TCG ATG TGT CTG CCG COT TTT	180
241	Ser Pro Thr Ser Cys Pro Pro Ile Cys Pro Gly Tyr Arg Thr Met Cys Leu Arg Asn Phe	240
301	S P T S C P P I C P O Y R W M C L R F	300
361	ATC ATA TTC CTC ATC ATC CTC GTG CTA TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT	360
421	Ile Ile Phe Leu Phe Ile Leu Leu Cys Leu Ile Phe Leu Val Leu Leu Asp Tyr	420
481	I I F L P I L L C L I F L L V L L D Y	480
541	CAA GGT ATG TTG CCG GTT TGT CTT CTA ATT CCA GGA TCA ACA ACC AGC ACG GGA CCG	540
601	Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Thr Thr Ser Thr Gly Pro	600
661	Q G M L P V C P L I P G S T T T S T O P	660
721	TGC AAA ACC TGC ACG AGT CCT GCT CAA GGA AAC TGT ATG TTT CCG TCC TGT TCC TGT ACA	720
781	Cys Lys Thr Cys Thr Thr Pro Ala Gln Gly Asn Ser Met Phe Pro Ser Cys Cys Thr	780
841	C K T C T T P A Q G N S M F P S C C T	840
901	AAA CCT ACG GAT GGA AAC TGC ACC TGT ATT CCG ATC CCA TCA TCC TGG GCT TTC GCA AAA	900
961	Lys Pro Thr Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Thr Ala Phe Ala Lys	960
1021	K P T D G N C T C I P I P S S W A F A E	1020
1081	TAC CTA TGG GAG TGG CCG TCA GTC GGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT	1080
1141	Tyr Leu Thr Gln Thr Val Arg Phe Ser Thr Thr Ser Leu Ser Leu Val Pro Phe Val	1140
1201	Y L W E W A S V R F S W L S L L V P V	1200
1261	CAA TGG TTC GTA GGG CTT TCC CCG ACT GTT TGG CTT TCA GCT ATA TGG ATG TGG TAT	1260
1321	Gln Thr Phe Val Gly Leu Ser Pro Thr Val Thr Leu Ser Ala Ile Thr Met Thr Tyr	1320
1381	Q W F V G L S P T V W L S A I W M W Y	1380
1441	TGG GGG CCA AGA CTG TAC TCC ATC GTT AGT CCG TTT ATC CCG GTG TTA CCA ATT TTC TTT	1440
1501	Trp Gly Pro Arg Leu Tyr Ser Ile Val Ser Pro Phe Ile Pro Leu Leu Pro Ile Phe Phe	1500
1561	W G P R L Y S I V S P P I P L L P I P F	1560
1621	TGT CTT TGG GTA TAC ATT	1620
1681	Cys Leu Thr Val Tyr Ile	1680
1741	C L W V Y I	1740

BEST AVAILABLE COPY

Fig. 3 Nucleotide sequence of the HBV surface antigen-encoding S gene of the HBV adw wild type
 (upper row from nt 1 to nt 678) as compared with the nucleotide sequence, which is sequenced
 from nt 127 to nt 588, of the novel variant HDB 05 (lower row, in which nucleotide
 differences are printed in bold type and bracketed when the mutations do not lead to any
 amino acid substitution)

1	ATG GAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTG CTC GTG TTA CAG GCG GGG TTT TTC	60
61	TTG TTG ACA AGA ATC CTC ACA ATA CCG CAG AGT CTA GAC TCG TGG ACT TCT CTC AAT	120
121	TTT CTA GGG GGA TCA CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC	180
	127: GGG GGA TCA CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC	
181	TCA CCA ACC TCC TGT CTT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG CCG CGT TTT	240
	TCA CCA ACC TCC TGT CTT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG CCG CGT TTT	
241	ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT	300
	ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT	
301	CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA ACA ACC AGG ACG GGA CCC	360
	CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA ACA ACC AGG ACG GGA CAA	
361	TGC AAA ACC TGC ACG ACT CCT GCT CAA GGA AAC TCT ATG TTT CCC TCC TGT TGC TGT ACA	420
	TGC AAA ACC TGC ACG ACT CCT GCT CAA (GGC) AAC TCT ATG TTT CCC (TCA) TGT TGC TGT ACA	
421	AAA CCT ACG GAT GGA AAC TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GCA AAA	480
	AAA CCT ACG GAT GGA AAC TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GCA AAA	
481	TAC CTA TGG GAG TGG GCC TCA GTC CGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT	540
	TAC CTA TGG GTG TGG GCC TCA GTC CGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT	
541	CAA TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG ATG ATG TGG TAT	600
	CGG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG 588	
601	TGG GGG CCA AGA CTG TAC TCC ATC GTT AGT CCC TTT ATC CCG CTG TTA CCA ATT TTC TTT	660
661	TGT CTT TGG GTA TAC ATT	678

BEST AVAILABLE COPY

Fig. 4 Nucleotide sequence of the S gene of the novel HBV variant HDB 05:

(nt 127 to nt 588) of the HBV surface antigen-encoding genome.

Only the nucleotide differences which lead to a change in the amino acid sequence are printed in bold.

```

- 127 GGG GGA TCA CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC 180
181 TCA CCA ACC TCC TGT CCT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT 240
241 ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT 300
301 CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA ACA AGA ACC AGT ACG GGA CAA 360
361 TGC AAA ACC TGC ACG ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT TGC TGT ACA 420
421 AAA CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA TTG TCC TGG GCT TTC GCA AAA 480
481 TAC CTA TGG GTG TGG GCC TCA GTC CGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT 540
541 CGG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG 588
```

Fig. 5 S gene nucleotide sequence (nt 127 to 588) and corresponding amino acid sequence (aa 43 to 196) of the novel HBV variant HDB 05 (amino acids which are substituted as compared with the HBV adw wild type are printed in bold and underlined)

127	GGG GGA TCA CCC GTG TGT CTT GGCCAA AAT TCG CAG TCC CCA ACCTCC AAT CAC	180
aa 43	G G S P V C L G Q N S Q S P T S N H	60
181	TCA CCA ACC TCC TGT OCT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT	240
61	S P T S C P P I C P G Y R W M C L R R F	80
241	ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT	300
81	I I F L F I L L L C L I F L L V L L D Y	100
301	CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA ACA AGA ACC AGT ACG GGA CAA	360
101	Q G M L P V C P L I P G S T R T S T G Q	120
361	TGC AAA ACC TGC ACG ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT TGC TGT ACA	420
121	C K T C T T P A Q G N S M F P S C C C T	140
421	AAA CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA TTG TCC TGG GCT TTC GCA AAA	480
141	K P T D G N C T C I P I P L S W A F A K	160
481	TAC CTA TGG GTG TGG GCC TCA GTC CGT TTC TCT TCG CTC AGT TTA CTA GTG CCA TTT GTT	540
161	Y L W Y W A S V R F S W L S L L V P F V	180
541	CGG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG	588
181	R W F V G L S P T V W L S A I W	aa 196

The following aa are substituted (x) in the HDB 05 variant as compared with the HBV adw wild type: T 115 (R), P 120 (Q), S 154 (L), E 164 (V) (all in the region of the a determinant) and Q 181 (R) (not in the region of the a determinant).

Fig. 6 Comparison of the amino acid sequences of the a determinant (aa 100 to aa 180) of the novel variant HDB 05 (lower row) and of the HBV adw wild type (upper row)

101	Q	G	M	L	P	V	C	P	L	I	P	G	S	T	T	T	S	T	G	P	120	
	Q	G	M	L	P	V	C	P	L	I	P	G	S	T	R	T	S	T	G	Q		
121	C	K	T	C	T	T	P	A	Q	G	N	S	M	F	P	S	C	C	C	T	140	
	C	K	T	C	T	T	P	A	Q	G	N	S	M	F	P	S	C	C	C	T		
141	K	P	T	D	G	N	C	T	C	I	P	I	P	S	S	W	A	F	A	K	160	
	K	P	T	D	G	N	C	T	C	I	P	I	P	L	S	W	A	F	A	K		
161	Y	L	W	E	W	A	S	V	R	R	F	S	W	L	S	L	L	V	P	F	V	180
	Y	L	W	Y	W	A	S	V	R	R	F	S	W	L	S	L	L	V	P	F	V	
181	Q	W	F	V	G	L	S	P	T	V	190											
	R	W	F	V	G	L	S	P	T	V												

The following aa are substituted (x) in the HDB 05 variant as compared with the HBV adw wild type:
 T 115 (R), P 120 (Q), S 154 (L), E 164 (V) - (all in the region of the a determinant) and Q 181 (R) (not in the region of the a determinant)

BEST AVAILABLE COPY